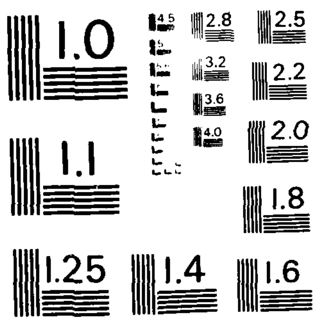
AD-A142 346 THEORY AND APPLICATIONS OF RANDOM FIELDS(U) ARMS CONTROL AND DISARMAMENT AGENCY WASHINGTON DC FIELD OPERATIONS DIV R J ADLER 06 FEB 84 AFOSR-TR-84-0458 1/1

UNCLASSIFIED AFOSR-83-0068 F/G 12/1 NL

END
DATE
FILMED
8 84
DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AFOSR-TR- 84-0458

AFOSR-83-0068

THEORY AND APPLICATIONS OF RANDOM FIELDS

Robert J. Adler
Faculty of Industrial Engineering & Management
Technion - Israel Institute of Technology
Technion City, Haifa, ISRAEL.

AD-A142 346

6 February 1984

Final Report, 1 January 1983 - 31 December 1983.

Approved for public release;
distribution unlimited.

Prepared for:

Technion Research & Development Foundation Ltd.
and
European Office of Aerospace Research & Development,
London, England.

84 06 19 031

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a SECURITY CLASSIFICATION AUTHORITY	3 DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited.
2b DECLASSIFICATION/DOWNGRADING SCHEDULE	
4 PERFORMING ORGANIZATION REPORT NUMBER(S)	5 MONITORING ORGANIZATION REPORT NUMBER(S) **AFOSR-TR- 84-0458**
6a NAME OF PERFORMING ORGANIZATION Technion - Israel Institute of Technology 6b. OFFICE SYMBOL (If applicable)	7a NAME OF MONITORING ORGANIZATION Air Force Office of Scientific Research
6c. ADDRESS (City, State and ZIP Code) Department of Industrial Engineering and Management, Technion City, Haifa, ISRAEL	7b. ADDRESS (City, State and ZIP Code) Directorate of Mathematical & Information Sciences, Bolling AFB DC 20332
8a NAME OF FUNDING/SPONSORING ORGANIZATION AFOSR 8b. OFFICE SYMBOL (If applicable) NM	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER AFOSR-83-0068
8c ADDRESS (City, State and ZIP Code) Bolling AFB DC 20332	10 SOURCE OF FUNDING NOS PROGRAM ELEMENT NO: 61102F PROJECT NO: 2304 TASK NO.: A5 WORK UNIT NO:
11 TITLE (Include Security Classification) THEORY AND APPLICATIONS OF RANDOM FIELDS	
12 PERSONAL AUTHOR(S) Robert J. Adler	
13a TYPE OF REPORT Final 13b. TIME COVERED FROM 1/1/83 TO 31/12/83	14 DATE OF REPORT (Yr., Mo., Day) FEB 84 15 PAGE COUNT 13
16 SUPPLEMENTARY NOTATION	
17 COSATI CODES FIELD / GROUP / SUB GR	18 SUBJECT TERMS (Continue on reverse if necessary and identify by block number) Random fields; chi-squared processes; suprema distributions; testing of normality; set indexed processes.

19 ABSTRACT (Continue on reverse if necessary and identify by block number)

The main results discussed in this final report include: (1) A discussion of the sample path behavior of chi-squared stochastic processes, with emphasis on the ways in which such processes exhibit non-normal behavior. (2) A discussion of the existence and path continuity of set indexed processes with independent increments. (3) Information on the distribution of suprema of random fields and empirical processes, with applications to stochastic modeling and statistical hypothesis testing.

20 DISTRIBUTION/AVAILABILITY OF ABSTRACT UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT ☐ DTIC USERS ☐	21 ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED	
22a. NAME OF RESPONSIBLE INDIVIDUAL Dr. Robert T. Smythe	22b TELEPHONE NUMBER (Include Area Code) (202) 767- 5028	22c OFFICE SYMBOL NM

DD FORM 1473, 83 APR EDITION OF 1 JAN 73 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

CONTENTS

0. Introduction

In the original research proposal for the grant being reported on, there were essentially three distinct, albeit related, projects. This report, for ease of writing, reading, and evaluation, is built around these projects in the following fashion: For each project I have included a complete recapitulation of what was presented in the original proposal. (A reader who is familiar with, and still remembers the details of, the original proposal can bypass the recapitulation.) Following this is, in each of the three cases, a report on the progress made towards realizing the goals of the proposal. The reports are generally rather brief, since they merely summarize results already presented in research papers, to which the reader can turn to for more details. Following each report is a brief comment on further research avenues (if any) opened up by work done to date.

Following the above three sections is a discussion of some work on empirical processes that was not originally foreseen in the proposal, but grew out of other projects in a natural way. This concludes the main body of the report. Lists of papers and conferences attended follow.

For the sake of completeness, we commence by including, directly from the original proposal, some background material on random fields.

1. Sample functions of random fields

Random fields are simply stochastic processes, $X(t)$, whose "time" parameter, t, varies over some Euclidean space, rather than over the real line. The simplest examples of these occur when the parameter space in question is two-dimensional, so that $X(t)$ is simply some sort of random surface. When the parameter space is three dimensional we have a field (such as ore concentration in a geological site) that varies over space, while when the dimension increases to four we are generally involved with space-time problems.

The last decade and a half has seen a large amount of scientific activity devoted to studying random fields - activity that has divided the subject into two basically distinct areas. The first covers problems in which the parameter t varies over a lattice, or similar, subset of Euclidean N-space. The models generally studied here, apart from being of intrinsic mathematical interest, are closely related to models of Statistical Mechanics such as the famed Ising model of magnetism. However, as interesting and important as these problems are, we shall, throughout this proposal, be concerned with the second class of problems - those that arise when the parameter is allowed to vary continuously over an appropriate region of N-space.

The theory of continuous parameter random fields is now quite substantial, and a large portion of it has recently been organised and coordinated in the monograph Adler (1). Roughly speaking, this theory can be summarised by breaking it into two distinct cases, as follows. In the first case, we assume that the sample functions (realisations) of the field satisfy certain regularity conditions, such as continuity, differentiability, etc. It is then possible to study such problems as the (statistical) distribution of the maximum value of the field, and the rate at which the field "crosses" (a term which requires careful definition) various levels. These problems turn out to be very important in applications of random fields to the study of rough surfaces,

as discussed in Section 2, below.

The second class of problems in the study of continuous parameter random fields arises when regularity assumptions such as those mentioned above are not imposed. Although the fields studied in this case can also be used to model natural phenomena, such as turbulence, geographical terrain and clusters of interstellar matter (see, esp. Mandelbrot (11)) they are primarily of mathematical interest, and are related to concepts such as local time and Hausdorff dimension. Since this class of random fields has only peripheral connection with the main subject of this proposal we shall say no more about it.

Although, as we have already mentioned, the theory of continuous parameter random fields with smooth sample functions is already quite substantial, it is important to note that in one sense at least it is still very restricted. This is a consequence of the fact that throughout the literature it is nearly always assumed that the random field being studied is Gaussian (normal), an assumption that has a substantial simplifying effect on the mathematics. There are two basic difficulties with such an assumption. The first, which comes from purely practical considerations, is that real life random fields, to which one might like to apply the theory, are often not Gaussian. For example, the rough metallic surfaces discussed in more detail in Section 2 are now known to be distinctly non-Gaussian (Adler (2)). Assuming, incorrectly, that they are Gaussian leads to the development of an unrealistic theory of surface structure - a theory that often fails to tie in with experiment. The second difficulty with the Gaussian assumption is that it hamstrings the Mathematician by limiting the phenomena available for his investigation to those that occur in this situation only.

In the following three sections we shall describe three projects in the study of random fields, ranging from the very applied to the purely theoretical. The common thread that runs through the three projects is the aim of extending both theory and applications of random fields by allowing for the use and appropriate development of non-Gaussian models.

2. Random field models of rough surfaces

It is now a well established fact that all surfaces used in engineering practice are rough when judged by the standards of molecular dimensions. This fact has played a major role in the development of the science and technology of Tribology, an area that, among other problems, is concerned with the nature of contact between surfaces under load and its relationship to problems such as friction, wear, and the conduction of heat and electricity between surfaces in contact.

Because of the difficulties inherent in observing what happens when two surfaces are forced together, Tribology has made substantial use of mathematical models. The basic idea underlying this has been to develop models of surface structure (at the microscopic level) and then apply these models, together with, for example, a theory of surface deformation, to predict observable (macroscopic) phenomena. Although substantial progress has been made in this area over the past fifteen years (see Archard et. al. (5) for a recent review) there is still very often disconcerting disagreement between theory and practice. This is despite the fact that very sophisticated random field models of rough surfaces have been employed.

It is the writer's sincere belief that this disagreement has a very simple explanation, and a reasonably simple remedy. Without exception, the engineering literature on rough surfaces models these random fields as being Gaussian, an assumption totally contra-indicated by almost all available data. This point was made recently in Adler and Firman (3), where both old and new data were presented and analysed to support this claim. Given, then, the non-Gaussian nature of true surfaces, it is not at all surprising that Gaussian models, regardless of their level of sophistication, fail to yield a theory that squares with practice.

The solution to this problem lies in the development of more realistic, necessarily non-Gaussian, models of rough surfaces. One such model, called the "inverted chi-squared" random field, has already been introduced (Adler and Firman (3)) and applied, in a much simplified form, and with surprising success, to the problem of determining the relationship between the true area of contact of

two rough surfaces and the force applied to bring them together (Adler (2)). The success of this simple application bodes well for further exploitation of this basic model. We propose to undertake such an exploitation in two distinct steps.

In Adler and Firman (3) the basic properties of inverted chi-squared random fields are derived in some detail. These properties include the height distribution of local maxima, the curvature distribution of maxima, etc., all of which form the basic building blocks of future application. The formulae that arose in that study are, however, extremely involved and awkward to manipulate. Thus, before any serious application of these fields can be made, it will be necessary to obtain simple approximations to these formulae that make further investigation feasible. This is the first step.

The second step involves applying the approximations to build inverted chi-squared models of rough surfaces. In principle, at least, this is reasonably straightforward, for all that really needs to be done is to replace the Gaussian assumption in existing models with the inverted chi-squared alternative. However, in practice, it can be expected that this will be anything but simple, for, along with substantial mathematical analysis, the non-Gaussian case can be expected to involve a substantial amount of numerical (computer) work to evaluate formulae that were often much simpler in the Gaussian case.

Finally, beyond the above application, there are a number of properties of inverted chi-squared random fields that are of substantial independent interest and that still remain to be investigated. For example, it is already known that these fields do not behave, in the neighbourhood of their local maxima, in anything like a Gaussian fashion. (See Adler (2) for details.) Thus it would be a most interesting, and undoubtedly worthwhile, task to investigate such aspects of the sample function behaviour of inverted chi-squared fields as their behaviour in the neighbourhood of high local maxima.

Thus, we have proposed here three sub-projects: the simplification of formulae describing local maxima, their application to real problems, and the more detailed behaviour of local maxima of chi-squared fields. We now describe

Progress

In late 1982 the general area of research outlined above was given to a doctoral student, Michael Aronowich, as a doctoral thesis problem. Since then, substantial progress, together with Aronowich, has been made. We began with the problem of the difficult formulae describing height distributions, etc., of local maxima. Here we soon realised that, unfortunately, many of the latter formulae in Adler and Firman (3) were incorrect, and so had to begin this project from scratch. Nevertheless, we were able to redo the calculations, and at the same time discover that when the "degree of freedom" parameter of a chi-squared process was odd, substantial simplification could be realised in the above formulae. The results of this work have been written up in a paper "Behaviour of χ^2 processes at extrema", recently submitted for publication.

Before turning to the problem of exploiting these results in an actual modelling application it seemed sensible to complete the theoretical study of chi-squared process behaviour at extrema by developing "Slepian model processes" for them. This turned out to be rather difficult, for although the basic theory here is precisely as in the Gaussian case, the computations and manipulations required in the chi-squared analogue were horrendous. Nevertheless, hard work and patience eventually yielded results, which will be described in a paper under preparation. Basically, the results obtained indicate that at high levels chi-squared maxima, after appropriate normalization, are not too different from their Gaussian counterparts. However, at low levels, extrema take a noticeably different form. Since low levels of chi-suared processes correspond to high levels of the "inverted" process, and it is the latter that is important for the modelling of rough surface behaviour, these results have potentially valuable applied significance. They also, of course, indicate successful completion of the third sub-project listed above.

Very little progress has been made in the application of these results to actual modelling problems, such as the modelling of rough surfaces. However the tools are now ready, and we hope to carry out such applications in the coming year under a continuation of the grant.

3. Differentiating between Gaussian and non-Gaussian fields

A very simple problem that arises as soon as one starts thinking about non-Gaussian stochastic processes and random fields is the following: Given some data, the realisation of a process or field, how does one determine if the data comes from a Gaussian or non-Gaussian model? Despite the simplicity of this question (and the ubiquity of Gaussian models) it is rather surprising that there is no satisfactory answer to it in the statistical literature other than that based on polyspectra.

The polyspectra approach to this problem (see, for example Brillinger (6)) has not been, we feel, very successful. The primary reason for this lies in the fact that the polyspectra procedure relies on testing for departure from Gaussianess in the cumulant structure of the process, and this is an aspect of any process that is singularly uninteresting and uninformative. Usually, the practicioner, and, indeed, the theoretician, is far more interested in the sample path structure of a process or field; that is, basic sample path properties, how often a process reaches a given level, and so forth.

Consequently, we propose a procedure for differentiating between Gaussian and non-Gaussian processes and fields by looking at their level crossing rates. (For information on these see, for processes, Cramér and Leadbetter (7), and for fields, Adler(1).) The basic idea here is that since Gaussian processes have a characteristic level crossing rate, and that under appropriate conditions numbers of level crossings are known to satisfy central limit theorems (e.g. Cuzick (8)) level crossing rates can be used as the basis of test statistics for Gaussianess.

Preliminary investigation shows this to be a powerful and efficient method for testing for normality. However a considerable amount of analytic work still remains to be done, as well as application of this method to various data sets and to simulations, before it can be used with confidence. Indeed, this problem is probably substantial enough in its own right to be the subject of an independent investigation. However, since it is intricately related with the problem discussed in Section 2, it is most likely best considered at the same time.

Progress

Of the main projects of the proposal, progress on this project has been the least satisfying, primarily because of personnel problems. Early in the year it became clear that although the idea of basing tests for Gaussianess on level crossings was a promising avenue, it was unlikely that exact statistical tests, which were powerful over a wide range of processes (i.e. covariance functions), could be developed. Thus the natural path to take was that of Monte Carlo testing; i.e. Given a realisation of some stochastic process [illegible] hypothesis Gaussian, use it to generate a number o[illegible] [illegible]aussian processes with the same parameters (i.e. covarianc[illegible] [illegible]unctions) and then compare the original realisation to the ([illegible] [illegible]onstruction Gaussian) simulations.

Such a test procedure relies, naturally, on quick and efficient computing procedures if it is to be useful. More importantly, in order to determine whether or not the procedure is statistically efficient, it is necessary in the early stages to conduct a massive simulation study of its properties.

During the year a variety of computing assistants developed a battery of computer programs to tackle the above task. This took much longer than it should have since good programmers tended to disappear quite quickly to better paying jobs outside of the Technion and poorer programmers...(The less said about these, the better.) In any case, by the end of the year programs to conduct a simulation study were ready, and some very early, preliminary, but promising, results obtained. I estimate that at this stage I am no more than two-three assistant months away from a set of nice results on testing for Gaussianess in one-parameter processes, and a corresponding paper.
Similar results for random fields, beacause of added computational problems, are probably another six-eight assistant months in the future.

4. Some theoretical problems

The problems described in the previous two sections were of an essentially applied nature, and the theoretical problems that were raised there arose as natural spin-offs of such investigations. However, it is envisaged that at the same time as the more applied aspects of this proposal will be carried out, fundamental, theoretical problems related to random fields will also be investigated. These, by their very nature, cannot be so well planned, for theoretical progress is always much more spasmodic than its applied counterpart. Nevertheless, it is envisaged that the following problems, at least, will be tackled.

It is often very easy to give the distribution of the maximum value of a stochastic process $X(t)$ as t varies over some interval on the line. (The classic case, $X(t)$ Brownian motion, is a textbook problem.) However, rather surprisingly, when we move to the problem of the distribution of the maximum of a random field it turns out that there is not a single non-trivial field for which this distribution is known. Goodman (13) has used some very elegant Banach space methods to obtain bounds on this distribution for the case $X(t)$ a multi-parameter Brownian motion, and it seems most likely that his approach can be generalised to cover a wider range of fields. We propose to use this approach, among others, to attack this problem.

A recent extension of the idea of random processes and fields has been to the notion of "set-indexed" processes, as discussed, for example, in Dudley (9,10). Here, the notion of entropy has been used to measure so-called "continuity classes" of sets for Gaussian processes. It would be of considerable interest to extend this theory from the Gaussian case to processes taking independent increments over disjoint sets. We propose to in fact carry out such an investigation, and construct and investigate such processes by combining the work of Pyke (12) on Gaussian set-indexed processes and Adler et.al. (4) on random fields with independent increments.

Progress

Progress on both of the above problems has been very pleasing. As regards the maximum problem, we succeeding in obtaining quite good upper and lower bounds to the probability

that the maximum of a two-parameter "pyramidal covariance function" field exceeds a given value. The field in question is a two-parameter version of the so-called "Slepian" or "triangular covariance" process on the line. This was the first time that such results have been given for stationary random fields, and the results will be published in the Annals of Probability in the near future.

The problem of the existence and continuity of set indexed processes with independent increments was also solved, in a joint paper with Professor Paul Feigin, also to be published in the Annals of Probability. We refer the reader to the paper itself for details.

An unexpected development during the year was a collaboration with Professor L. D. Brown of Cornell University, which resulted in the solution of a very old problem on empirical processes, described below.

5. Empirical processes.

Since none of the empirical process work has yet been written up, we shall describe it in some detail. The problem is as follows:

Let $X_1, X_2, \ldots, X_n$ be a sequence of i.i.d. observations from some k-dimensional distribution function F, and let F_n be the multi-parameter empirical distribution function based on the X_i. Then, under very weak assumptions on F, $n^{\frac{1}{2}}(F_n - F)$ converges weakly to a limit process, say W_F, that is a multi-parameter generalisation of the Brownian bridge. For statistical hypothesis testing, a statistic of interest is the Kolmogorov-Smirnov statistic $\sup(F_n(x) - F(x))$, whose distribution, asymptotically, can be obtained from the distribution of $\sup(W_F(x))$. In the multi-parameter situation under consideration the distribution of the Kolmogorov-Smirnov statistic depends on F so that, unlike the one-dimensional situation, it is not distribution free. However, the situation is even worse than this, since except for the case of F uniform on the unit square, nothing was known about the distribution of $\sup(W_F(x))$.

What Brown and I managed to show was the following: There exist constants $c(F,k)$ and $C(k)$ such that for all $\lambda>0$

$$c(F,k)\lambda^{2(k-1)}e^{-2\lambda^2} \leqq P\{\sup_x W_F(x) > \lambda\} \leqq C(k)\lambda^{2(k-1)}e^{-2\lambda^2}.$$

When $k=2$ the constants can be more or less identified, thus permitting the development of Kolmogorov-Smirnov tests for bivariate problems. For general k the bound can be used to obtain sharp upper-lower class results for the growth of $\sup (F_n(x) - F(x))$ with n.

These results are currently being written up in two papers; one with the bounds above, their proofs, and the upper-lower class theorems, and one explaining the significance of these results to statisticians.

A rather surprising and pleasing off-shoot of this work is that it seems that it will be possible to apply the methodology of the proofs to study the tail behaviour of the suprema of a wide class of Gaussian processes, and to remove the ε in the well-known Fernique-Landau-Marcus-Shepp result that for large enough λ and all $\varepsilon>0$,

$$\log(P\{\sup_t X(t)>\lambda\}) < \lambda^2\{ \varepsilon - \tfrac{1}{2}(\sup_t \mathrm{var}X(t))^{-1} \}.$$

This work is still in progress.

References

1. Adler R.J. 1981 The Geometry of Random Fields Wiley, London.
2. Adler R.J. 1981 Random field models in surface science Bull. Int. Statist. Inst. in press.
3. Adler R.J. & Firman D. 1981 A non-Gaussian model for random surfaces Philos. Trans. Roy. Soc. in press.
4. Adler R.J., Monrad D., Scissors R. & Wilson R.J. 1982 Representations, decompositions, and sample function continuity of random fields with independent increments Stoch. Proc. Appl. in press
5. Archard J.F., Hunt R.T. & Onions R.A. 1975 Stylus profilometry and the analysis of the contact of rough surfaces. In The Mechanics of Contact Between Deformable Bodies (eds. de Pater & Kalker) Delft Univ. Press, Belgium.
6. Brillinger D. 1975 Time Series, Data Analysis and Theory Holt, Rinehart & Winston, New York.
7. Cramér H. & Leadbetter M.R. Stationary and Related Stochastic Processes Wiley, New York.
8. Cuzick J. 1976 A central limit theorem for the number of zeros of a stationary Gaussian process Ann. Probability 4 547-556.
9. Dudley R.M. 1973 Sample functions of the Gaussian process Ann. Probability 1 66-103.
10. Dudley R.M. 1976 Central limit theorems for empirical measures Ann. Probability 6 899-929.
11. Mandelbrot B. 1977 Fractals: Form, Chance, and Dimension Freeman, San Francisco.
12. Pyke R. 1977 The Haar function construction of Brownian motion indexed by sets Technical Report #35 Mathematics Department, University of Washington.
13. Goodman V. 1976 Distribution estimates for functionals of the two parameter Wiener process Ann. Probability 4 977-982.

6. Publications prepared under the grant

1. Random fields, Invited survey article to appear in Encyclopedia of Statistical Sciences, eds. N. Johnson and S. Kotz, Wiley, 1984.
2. The supremum of a particular Gaussian field, accepted for publication in Annals of Probability, 1984.
3. On the cadlaguity of random measures, (joint with P. D. Feigin) accepted for publication in Annals of Probability, 1984.
4. Behaviour of χ^2 processes at extrema, (joint with A. Aronowich), submitted.
5. Tail behaviour for the suprema of empirical processes, (joint with L. D. Brown), in preparation.
6. Multivariate Kolmogorov-Smirnov tests, (joint with L. D. Brown), in preparation.

7. Conferences attended & visits

During the summer of 1983 I visited the following American universities, to work with the following colleagues:

Cornell; Brown, Prabhu, Taqqu; 3 weeks.

MIT; Dudley; 3 days

University of Massachusetts, Amherst; Horowitz, Geman, Rosen, 3 days.

I also attended the following conferences, at which I presented the following papers:

12th Conference for Stochastic Processes & Their Applications, Ithaca, July 11-15, "Exact distributions of the maximum of some Gaussian random fields".

185th IMS Meeting, Toronto, August 15-18, "Some sample path properties of random fields" (Invited paper).

MED

8